UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 05-62
Economics Series 31
November 2005

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624 98 75

THE AXIOMATIC PROPERTIES OF AN ENTROPY BASED INDEX OF SEGREGATION*

***Ricardo Mora* [1] *and Javier Ruiz-Castillo* [2]**

Abstract

This paper reviews the properties suggested in the methodological literature on the measurement of occupational gender segregation. It is found that an index of (relative) segregation based on the entropy concept, I_E, satisfies thirteen basic axioms previously proposed in the single-dimensional case, and can be expressed as the sum of a between-group and a within-group term both for any partition of the set of occupations and in the two-dimensional case.

Keywords: gender segregation measurement; axiomatic properties.

* **Acknowledgements:**.The authors acknowledge financial support from the Spanish DGI, Grants BEC2003-03943 and SEJ2004-01959.

[1] Ricardo Mora, Universidad Carlos III de Madrid. Departamento de Economía. C/ Madrid 126, Getafe 28903, Madrid, Spain. E-mail: ricmora@eco.uc3m.es
[2] Javier Ruiz- Castillo. Universidad Carlos III de Madrid. Departamento de Economía . C/ Madrid , 126,. Getafe 28903 Madrid. E-mail: jrc@eco.uc3m.es

1. INTRODUCTION

Social scientists have long been interested in the problem of segregation in the labor market by gender, that is, the tendency of men and women in the employment population to be differently distributed across occupations.[3] The information contained in the joint distribution of gender and occupation is usually summarized by means of numerical indices of segregation. In spite of the large volume of contributions, most of the proposed indices fall into the following three categories.

The first family of indices refers to those inspired by the Index of Dissimilarity, ID, first proposed in Duncan and Duncan (1955). The popularity of this index is based on its appealing interpretation as the proportion of male or female workers that would have to be removed without replacement in order to make every occupation contain the same gender mix exhibited by the labour force as a whole. This interpretation is at the core of the development of several variants of the index.[4] A second approach exploits the connection between the measurement of income inequality and the measurement of gender segregation viewed as the inequality in the distribution of the employed population across occupations. This is the case of indices inspired in the Gini index of income inequality, as well as the family of Atkinson's indices, the coefficient of variation, the so-called square root index, or one of Theil's measures.[5] Finally, a statistical approach to gender segregation measurement has been recently advocated

[3] The seminal article on (residential) segregation is Duncan and Duncan (1955). For recent contributions to gender segregation, see the special issues of the *Journal of Econometrics*, 1994, **61**(1), and *Demography*, 1998, **35**(4), as well as the treatise by Flückiger and Silber (1999).

[4] See Cortese *et al.* (1976), Moir and Selby Smith (1979), Lewis (1982), Karmel and MacLachlan (1988), Silber (1992), and Watts (1992). The index and its variants have become so dominant after the "index wars" (Peach, 1975), that concern has recently been voiced about a situation in which it is generally "assumed that sex segregation is simply whatever ID measures" (Grusky and Charles, 1998).

under the argument that the conventional practice of using a scalar index to describe gender segregation differences over time and/or across countries must be embedded in a testable model. This is the case of Charles (1992, 1998), Charles and Grusky (1995) and Grusky and Charles (1998), who propose a log-multiplicative model, or Kakwani (1994) who develops a procedure based on the *F*-distribution to test whether gender segregation has increased or decreased significantly within any two periods or across any two countries.

This paper defends the use of an index, I_E, which is based on the entropy concept used in information theory and has a rather simple and nice statistical interpretation. It was first introduced in the segregation literature by Theil and Finizza (1971) and Fuchs (1975), and has recently been extended by Herranz *et al.* (2005) and Mora and Ruiz-Castillo (2003a, 2004) in a series of papers that exploit its additive decomposability properties both when the set of occupations is partitioned into a number of non-overlapping subgroups, as well as when segregation takes place along two or more dimensions.

Naturally, two segregation indices may show different trends in a given country, and may produce different country rankings in international comparisons.[6] Thus, the design of measures with desirable properties is a central methodological issue, and the merits of competing indices are regularly debated.[7] For our purposes, the properties of segregation indices discussed in the literature can be classified into four groups. First, there are a number

[5] See, *inter alia*, Duncan and Duncan (1955), Schwartz and Winship (1979), Butler (1987), Silber (1989a, 1989b), Hutchens (1991, 2001, 2004), and Flückiger and Silber (1999).

[6] For some evidence in this respect, see *inter alia* Jonung (1984), James and Taeuber (1985), Karmel and MacLachlan (1988), Blackburn *et al.* (1993), Anker (1998), and Flückiger and Silber (1999).

[7] See *inter alia*, the methodological contributions by James and Taeuber (1985), Siltanen (1990), Hutchens (1991, 2001, 2004), Watts (1992, 1997, 1998a, 1998b), Blackburn *et al.* (1993, 1995), Kakwani (1994), Charles (1992), Charles and Grusky (1995), Grusky and Charles (1998), and Flückiger and Silber (1999).

of basic desirable characteristics for the case in which gender segregation takes place along a single dimension, say occupation. These properties might be required from both relative as well as absolute segregation measures, although our discussion is restricted to relative segregation indices. Second, for all possible partitions of the set of occupations it is useful that overall segregation can be expressed as the sum of two terms. The first term captures the weighted sum of the segregation *within* each subgroup of occupations, while the second term measures the *between-group* segregation computed as if every occupation had the mean number of males and females of the occupational subgroup to which it belongs.[8] Third, since segregation measures are usually computed using sample observations, an additional desirable property for a measure of segregation is that it is embedded in a statistical framework that permits the testing of hypothesis on gender segregation in occupations. Fourth, there is an important group of invariance axioms that are motivated by the interest of making intertemporal and international comparisons of segregation levels and serve to make precise what is meant by a margin-free index, that is, a segregation index that is independent from changes in the overall share of employment by gender (composition invariance), and from changes in the occupational structure (occupational invariance).

Regarding the I_E index, it has been shown elsewhere that it is embedded in a statistical framework and that even though it is not margin free, it can be decomposed to isolate margin free terms; under certain conditions, these margin-free terms are also shown to have a simple intuitive statistical interpretation (see Mora and Ruiz-Castillo, 2005). The purpose of this paper

[8] Similarly, when segregation takes place along two dimensions, say educational level and occupation, it is useful that overall segregation can be decomposed into a term that captures the between-group segregation induced by

is twofold. First, it reviews recent methodological contributions to the measurement of gender segregation and its alternative notions. Second, it is shown that the entropy based index I_E satisfies thirteen basic properties in the single-dimensional case, and that it is decomposable into a *between-group* and a *within-group* term both for any partition of the set of occupations and in the two-dimensional case.

The rest of the paper contains three sections. Section 2 reviews the main axioms discussed in the literature. Section 3 presents the I_E index of segregation and proofs that it satisfies those axioms. Section 4 briefly reviews other well-known segregation measures in the light of the axioms discussed in the paper and offers some concluding comments.

2. BASIC AXIOMS

2. 1. The Single-dimensional Case. Notation

Assume an economy with *J* occupations, indexed by $j = 1,\ldots, J$. The usual data available in empirical situations can be organized into the following (3 x (J + 1)) array

$$\begin{pmatrix} F_1 & F_2 & \cdots & F_J & F \\ M_1 & M_2 & \cdots & M_J & M \\ T_1 & T_2 & \cdots & T_J & T \end{pmatrix} = \begin{pmatrix} \mathbf{f} & F \\ \mathbf{m} & M \\ \mathbf{t} & T \end{pmatrix}$$

where $\mathbf{f} = (F_1, F_2,..., F_J)$, $\mathbf{m} = (M_1, M_2,..., M_J)$ and $\mathbf{t} = (T_1, T_2,..., T_J) = (F_1 + M_1, F_2 + M_2,..., F_J + M_J)$ are the (1 x *J*) vectors of females, males, and people, respectively, employed in each occupation, whereas $F = \Sigma_j F_j$, $M = \Sigma_j M_j$ and $T = \Sigma_j T_j$ are, respectively, the total number of

one of the classification variables, and a second term that measures the segregation induced by the second variable within the subgroups defined by the first one.

females, males, and people in the economy.

For later reference, define three types of (1 x J) vectors. First, the vectors $\mathbf{s}^{\mathrm{f}} = (s_{f1},\ldots, s_{fJ}) = (F_1/F,\ldots, F_J/F)$, $\mathbf{s}^{\mathrm{m}} = (s_{m1},\ldots, s_{mJ}) = (M_1/M,\ldots, M_J/M)$ and $\mathbf{s}^{\mathrm{t}} = (s_{t1},\ldots, s_{tJ}) = (T_1/T,\ldots, T_J/T)$, capturing the frequency distributions over occupations of females, males and people, respectively. Second, the vectors $\mathbf{w} = (w_1,\ldots, w_J) = (F_1/T_1,\ldots, F_J/T_J)$ and $(\mathbf{1} - \mathbf{w}) = (1 - w_1,\ldots, 1 - w_J) = (M_1/\mathrm{T}_1,\ldots, M_J/T_J)$ of female and male employment shares in all occupations. Third, the vector of gender ratios $\mathbf{r} = (r_1,\ldots, r_J) = (F_1/M_1,\ldots, F_J/M_J)$. Finally, denote the overall female and male shares by $W = F/T$ and $(1 - W) = M/T$, respectively, and the overall gender ratio by $R = F/M$.

In many contexts, numerical indices serve to summarize the degree of gender segregation prevailing in the entire economy, and provide a concise means of presenting the dominant trends that may be hidden in a detailed occupation by occupation study. For the sake of generality, a distribution of people across gender and occupations will be identified in the sequel by a 6-tuple $(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T)$. Any scalar index of segregation, θ, can then be seen as a unique real non-negative valued and continuous function of $(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T)$, $\theta = \theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T)$.[9]

2. 2. Thirteen Basic Axioms

Among others, James and Taeuber (1985), Siltanen (1990), Kakwani (1994), and Hutchens (1991, 2001) have proposed a number of desirable properties for an index of segregation. These properties will be presented below as axioms. However, these axioms need not be considered all desirable at the same time. As in Kakwani (1994), the purpose here is not so much to justify them as to provide a framework for comparing various segregation indices.[10]

All notions of occupational gender segregation stem from an idea of association between gender and occupational category. As indicated in the Introduction, it is usually understood that an index of gender segregation $\theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T)$ measures the extent to which the female and the male distributions differ across occupations. This is why some of the basic axioms presented in the sequel (in particular, A.1, and A.6 to A.9), as well as definition 1 will be couched in terms of the vectors $\boldsymbol{s}^{\mathrm{f}}$ and $\boldsymbol{s}^{\mathrm{m}}$.

In the literature on income inequality, it is customary to distinguish between indices that focus on income differences and indices that focus on income shares (see Kolm, 1999). In the first case, the measure of income inequality is invariant to equal additions to all incomes (translation invariance), and indices are referred to as absolute indices. In the second case, income inequality is not affected by proportional changes in all incomes (scale invariance), and indices are referred to as relative indices. Scale and translation invariance correspond to two particular inequality views so that the choice among them is normative and depends on value judgements. In the segregation literature, most indices entail a relative view in which relative magnitudes are all that matters. Formally:

---

[9] Of course, this formal framework is equally well suited for the measurement of other segregation phenomena, such as the segregation exhibited by the distribution of black and white students over schools in a given school district.

[10] This approach can be contrasted to Hutchens (2004) and Chakravarty and Silber (1992), the only two studies in the segregation literature that attempt an axiomatic characterization of specific numerical measures.

Axiom 1: (*Size Invariance*, James and Taeuber, 1985) Let ($\mathbf{f}'$, F', $\mathbf{m}'$, M', $\mathbf{t}'$, T') = ($\lambda\mathbf{f}$, λF, $\lambda\mathbf{m}$, λM, $\lambda\mathbf{t}$, λT) where λ is a positive scalar. Then $\theta(\mathbf{f}', F', \mathbf{m}', M', \mathbf{t}', T') = \theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T)$. Clearly, under A.1, all relative magnitudes –namely, s^f, s^m, s^t, $\mathbf{w}$, $(1 - \mathbf{w})$, $\mathbf{r}$, W, (1 – W), and R- remain constant.[11] □

Explicit in the calculation of any index is the specification of two counterfactual distributions that capture the ideas of complete integration and complete segregation. Within the above notion of occupational gender segregation, there is broad agreement on the meaning of what these two distributions should be.

Axiom 2: (*Complete Integration*, Kakwani 1994) Let ($\mathbf{f}$, F, $\mathbf{m}$, M, $\mathbf{t}$, T) be such that $\mathbf{s}^f = \mathbf{s}^m$. Then $\theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T) = 0$. □

Notice that this relative notion of complete integration is not the only one. In an absolute context, Chakravarty and Silber (1992) suggest stronger notion of complete integration, according to which there is no gender segregation if and only if $F_j = M_j$ for all j.

Axiom 3: (C*omplete Segregation*, Kakwani 1994) Let ($\mathbf{f}$, F, $\mathbf{m}$, M, $\mathbf{t}$, T) be so that F_j (M_j) > 0 implies M_j (F_j) = 0 for all j. Then $\theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T) = 1$. □

This axiom implies that the index should have a maximum value of unity when females and males are in separate occupations.

The next two axioms capture two different symmetry notions.

Axiom 4: (*Symmetry in Groups*, Kakwani 1994 and Hutchens 1991) Let $\mathbf{f}'$ and $\mathbf{m}'$ be two permutations of $\mathbf{f}$ and $\mathbf{m}$, respectively. Then $\theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T) = \theta(\mathbf{f}', F, \mathbf{m}', M, \mathbf{t}', T)$. □

Axiom 5: (*Symmetry in Types*, Kakwani 1994 and Hutchens 2001) $\theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T) = \theta(\mathbf{m}, M, \mathbf{f}, F, \mathbf{t}, T)$. □

That a segregation index should be insensitive to whether men or women are labeled as "males" or "females" is a reasonable value judgment. However, Hutchens (2004) forcefully argues that, as long as it implies that movements across groups of people and income are equivalent, A.5 is less compelling for a measure of income inequality.

For the next axioms, it is useful to introduce the following:

Definition 1: An occupation j is *female dominated* if and only if $s_{fj} > s_{mj}$. □

Axiom 6: (*Weak Principle of Transfers*, James and Taeuber, 1985, Kakwani 1994) If there is a small shift of the female (male) labor force from a female- (male-) dominated occupation to a male- (female-) dominated occupation, the segregation index must decrease. □

Siltanen (1990) and Watts (1992) propose a somewhat stronger condition than A.6, which is also closely related to the following:

Axiom 7: (*Movement between Groups*, Hutchens 1991) Let $M'_h = M_h = M'_j = M_j$ for any h, j. Assume that there are two occupations i and k such that: (a) $(s_{fi}/s_{mi}) < (s_{fk}/s_{mk})$, (b) $F'_i = F_i$ - d and $F'_k = F_k$ + d, for $0 < \mathrm{d} \le F_i$, and (c) $F'_j = F_j$ for any $j \neq i, k$. Then $\theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T) < \theta(\mathbf{f}', F, \mathbf{m}', M, \mathbf{t}, T)$. □

This disequalizing movement is similar to a regressive transfer in the income inequality literature. It reduces the presence of women in a given occupation, and it increases it in an occupation that originally has a higher ratio of women to men. Therefore, A.7 plays here the

[11] For a study that focuses on translation invariant segregation indices that represent an absolute view of segregation, see Chakravarty and Silber (1992).

same role as the Pigou-Dalton principle in the income inequality literature.

Kakwani (1994) argues that a segregation index must be sensitive to any shift in the labor force from one occupation to another. The two previous axioms refer to shifts from a female (male) to a male (female) dominated occupation. In order to determine the sign of the change in the index when the shift takes place between two female (or two male) occupations new value judgments are introduced in the next two axioms.

Axiom 8: (Kakwani 1994) If i and k are both female (male) dominated occupations with exactly equal gaps, $|s_{fi} - s_{mi}| = |s_{fk} - s_{mk}|$, then a small shift of the female (male) labor force from occupation i to k should reduce (increase) the segregation index whenever $s_{ti} < s_{tk}$ ($s_{ti} > s_{tk}$). □

Axiom A.8 represents a strong value judgment implying that, in a pair of female (male) occupations, it is more desirable to increase (reduce) the male-female ratio in the smaller one. The justification offered by Kakwani (1994) is that the relative importance of an occupation is inversely related to the probability that a person belongs to it, that is, it is inversely related to its size. This is reflected in the fact that small occupations are generally among the higher paid ones. Therefore, gaps among them should be given larger weights.

On the other hand, whenever the two occupations have the same size, the next axiom requires that a small shift in the labor force from one occupation to another should reduce the segregation index if the gap between the female and the male employment proportions is larger in the first one.

Axiom 9: (Kakwani 1994) If i and k are both female- (male-) dominated occupations with size $T_i = T_k$, then a small shift of the female (male) labor force from occupation i to k

should reduce (increase) the segregation index if $|s_{fi} - s_{mi}| > |s_{fk} - s_{mk}|$ $(|s_{fi} - s_{mi}| < |s_{fk} - s_{mk}|)$. □

In the context of residential segregation, Zoloth (1976) introduced the notion of *diminishing payoffs to desegregation* as a useful property from a policy point of view, arguing that the cost of additional desegregation rises with the level of desegregation already achieved. This notion is analogous to the property of *decreasing returns of inequality in proximity* in Kolm (1999), or the *transfer sensitivity* property in Shorrocks and Foster (1987) in the income inequality literature. This idea can be formulated as a stronger condition than A.7:

Axiom 10: (*Increasing Returns to a Movement Between Groups*, Zoloth 1976) Let $M''_h = M'_h = M_h = M''_j = M'_j = M_j$ for any h, j. Assume that there are two occupations i and k such that: (a) $(s_{fi}/s_{mi}) < (s_{fk}/s_{mk})$, (b) $F''_i = F'_i - \mathrm{d}$, $F''_k = F'_k + \mathrm{d}$, $F'_i = F_i - \mathrm{d}$ and $F'_k = F_k + \mathrm{d}$, for $0 < 2\mathrm{d} \leq F_i$, and (c) $F''_j = F'_j = F_j$ for any $j \neq i, k$. Then $[\theta(\mathbf{f}'', F, \mathbf{m}'', M, \mathbf{t}, T) - \theta(\mathbf{f}', F, \mathbf{m}', M, \mathbf{t}, T)] > [\theta(\mathbf{f}', F, \mathbf{m}', M, \mathbf{t}, T) - \theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T)] > 0$. □

Several contributions in the literature have emphasized the importance of basic aggregation properties. In this context, the simplest requirement that an index of segregation must satisfy is that a group with no members should have no effect on segregation. Consequently, one can delete occupations that contain no people without affecting measured segregation.

Axiom 11: (*Zero Member Independence*, Hutchens 2001). Let $(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T)$ and $(\mathbf{f}', F, \mathbf{m}', M, \mathbf{t}', T)$ be identical except that $(\mathbf{f}', F, \mathbf{m}', M, \mathbf{t}', T)$ includes an occupation $J + 1$ with no members, $T_{J+1} = 0$, that is excluded from $(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T)$. Then $\theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T) = \theta(\mathbf{f}', F, \mathbf{m}',$

M, $\mathbf{t}'$, T). □

For the next property, it is useful to introduce the notion of a proportional division, an operation that divides an existing occupation into several new ones so that the gender ratio of female to male workers in the new occupations remains equal to the original (predivision) ratio.

Definition 2: (Hutchens 2001) Let N be an integer. A distribution ($\mathbf{f}'$, F, $\mathbf{m}'$, M, $\mathbf{t}'$, T) is said to be obtained from ($\mathbf{f}$, F, $\mathbf{m}$, M, $\mathbf{t}$, T) through *a proportional division* of, say, occupation J, into $N + 1$ new ones, if $F'_j = F_j$ and $M'_j = M_j$ for all $j \neq J$, and $F'_i = F_i/(N + 1)$ and $M'_i = M_i/(N + 1)$, so that $r'_i = r_i$ for all $i = J, J + 1, \ldots, J + N$. □

The next axiom requires that an index be unaffected by the division of an occupation into units with identical segregation patterns. As pointed out by James and Taeuber (1985), this principle has no analogue in the literature on income inequality measurement. It allows the comparison of economies with a different number of occupations by artificially equalizing those numbers with the help of a suitable division or combination of occupations.

Axiom 12: (*Organizational Equivalence*, James and Taeuber 1985, or *Insensitivity to Proportional Divisions*, Hutchens 2001) Let ($\mathbf{f}'$, F, $\mathbf{m}'$, M, $\mathbf{t}'$, T) be obtained from a proportional division of an occupation of ($\mathbf{f}$, F, $\mathbf{m}$, M, $\mathbf{t}$, T). Then $\theta(\mathbf{f}', F, \mathbf{m}', M, \mathbf{t}', T) = \theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T)$. □

Finally, in many contexts we are interested not only in the extent of gender segregation, but also in the actual pattern that characterizes this phenomenon in each occupation. Similarly, it may be useful to measure the contribution of each occupation to overall gender segregation. As long as a notion of local segregation is introduced, further requirements on the relation between overall and local measures might be appropriate. Suppose, for instance, that after a rearrangement of the population segregation rises in each occupation. It then seems reasonable to require that the overall segregation value does not decrease. To formalize these ideas, assume that the relevant information about gender segregation in each occupation j can be described by the 6-tuple (F_j, F, M_j, M, T_j, T) where, as before, $F = \Sigma_j F_j$, $M = \Sigma_j M_j$ and $T = \Sigma_j T_j$. A local index of gender segregation in that occupation, θ_j, will be a real valued and continuous function $\theta_j = \theta_j(F_j, F, M_j, M, T_j, T)$ that it is bounded and satisfies A.4. Now it is possible to state the following strong requirement:

Axiom 13: (*Additivity*) The segregation index θ is said to be additive if there exists a non-decreasing and continuous real valued function F such that, for any $(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T)$, $\theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T) = F\{\Sigma_j \theta_j(F_j, F, M_j, M, T_j, T)\}$. □

The notion of segregation used so far refers to a situation in which the vectors $\boldsymbol{s}^{\mathrm{f}}$ and $\boldsymbol{s}^{\mathrm{m}}$ are different. However, segregation can also be said to exist (i) when the female shares w_j differ across occupations, as in Anker (1998)'s measure of gender dominated occupations and the entropy measure first proposed by Theil and Finizza (1971), or (ii) when it is the gender ratios r_j that differ across occupations, as in the index first suggested in Charles (1992). Since $w_j \neq w_k$ for any $j, k \in \{1,\ldots, J\}$ if and only if $r_j \neq r_k$, these two notions need not be treated separately.[12] In any case, all axioms presented in terms of the vectors $\boldsymbol{s}^{\mathrm{f}}$ and $\boldsymbol{s}^{\mathrm{m}}$ (A.2, and A.7 to

[12] This is true under the assumption that there is some positive male and female employment in each occupation. Otherwise, gender ratios are not well defined.

A.10), as well as Definition 1 can be equivalently written in terms of the vector(s) $\boldsymbol{w}$ (or $\boldsymbol{r}$).[13]

2. 3. Decomposability Properties

A. The Case of a Partition of the Set of Occupations[14]

Consider an island A with J occupations, indexed by $j = 1,\ldots, J$, and an island B with a different set of K occupations indexed by $k = J + 1,\ldots, J + K$. Assume that in island A the total number of females and males, F^A and M^A, respectively, are uniformly distributed across the J occupations, so that $F_j = F^A/J$ and $M_j = M^A/J$ for all j. In this case, since $s_{fj} = s_{mj} = 1/J$ for all j, there is no segregation in island A. Similarly, assume that in island B the total number of females and males, F^B and M^B, respectively, are uniformly distributed across the K occupations, so that $F_k = F^B/K$ and $M_k = M^B/K$ for all k. Again, since $s_{fk} = s_{mk} = 1/K$ for all k, there is no segregation in island B. Now assume that the two islands form a confederation. In spite of the fact that there is no segregation within the two islands, as long as $F^A/(F^A + F^B)$ is different from $M^A/(M^A + M^B)$ –in which case we will also have that $F^B/(F^A + F^B)$ is different from $M^B/(M^A + M^B)$- there will be some segregation in the confederation as a whole. As in the income inequality literature, this example suggests the usefulness of being able to decompose overall segregation in the confederation into a within-island and a between-island component.

More generally, assume that the set of J occupations is partitioned into I groups,

[13] Note that $s_{fj} > s_{mj}$ if and only if $w_j > W$. However, if $w_j = k\,W$, $k \in (0, 1/W)$, then $s_{fj} = f(k,W)\, s_{mj}$ where $f(k, W) = [\,(1/kW) - 1]^{-1}$. Thus, the correspondence between the two notions of "dominance" is a non-linear monotonic

indexed by $i = 1,\ldots, I$, and denote by G_i the number of occupations in group i, so that $\Sigma_i G_i = J$. Let F_{ij}, M_{ij} and $T_{ij} = F_{ij} + M_{ij}$ be the number of females, males, and people, respectively, in occupation j within group i; let $F_i = \Sigma_{j\in G_i} F_{ij}$, $M_i = \Sigma_{j\in G_i} M_{ij}$ and $T_i = \Sigma_{j\in G_i} T_{ij}$ be the total number of females, males and people in group i, and let $\mathbf{f}^i = (F_{i1}, F_{i2},\ldots, F_{iG_i})$, $\mathbf{m}^i = (M_{i1}, M_{i2},\ldots, M_{iG_i})$, and $\mathbf{t}^i = (T_{i1}, T_{i2},\ldots, T_{iG_i})$ be, respectively, the gender and people's frequencies across the G_i occupations in group i. Let $F = \Sigma_i F_i$, $M = \Sigma_i M_i$ and $T = \Sigma_i T_i$ be the overall number of females, males and people, respectively. The distributions of F, M, and T across the J occupations in the economy as a whole can then be written as $\mathbf{f} = (\mathbf{f}^1,\ldots, \mathbf{f}^I)$, $\mathbf{m} = (\mathbf{m}^1,\ldots, \mathbf{m}^I)$, and $\mathbf{t} = (\mathbf{t}^1,\ldots, \mathbf{t}^I)$, respectively.

Several measures of segregation are then available in this situation: (i) an overall measure of segregation, $\theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T)$; (ii) a *within-group* measure of segregation $\theta^i(\mathbf{f}^i, F_i, \mathbf{m}^i, M_i, \mathbf{t}^i, T_i)$ for each i; and (iii) a *between-group* measure of segregation computed as if every occupation j had the mean number of males and females of the group i to which it belongs. Thus, the between-group segregation measure is defined as $\theta(\mathbf{f}', F, \mathbf{m}', M, \mathbf{t}', T)$, where $\mathbf{f}' = \{(F_1/G_1)\mathbf{e}^{G_1},\ldots, (F_I/G_I)\mathbf{e}^{G_I}\}$, $\mathbf{m}' = \{(M_1/G_1)\mathbf{e}^{G_1},\ldots, (M_I/G_I)\mathbf{e}^{G_I}\}$, $\mathbf{t}' = \{(T_1/G_1)\mathbf{e}^{G_1},\ldots, (T_I/G_I)\mathbf{e}^{G_I}\}$ and, for each i, $\mathbf{e}^{G_i}$ is a G_i-dimensional vector of ones. In this context, a convenient property is that the overall measure of gender segregation can be expressed as the sum of two

increasing function of k and W. It is then possible to think of situations whereby a change in k is offset in $f(k,W)$

components: a *between-group* term, which captures the gender segregation at the higher (group) level of aggregation; plus a weighted sum of *within-group* terms, where each of them captures the occupational gender segregation induced within each group.[15]

Axiom 14: (*Additive Decomposability*) There exist $\nu_i \geq 0$ for all i with $\Sigma_i \nu_i = 1$, so that $\theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T) = \Sigma_i \nu_i \theta^i(\mathbf{f}^i, F_i, \mathbf{m}^i, M_i, \mathbf{t}^i, T_i) + \theta(\mathbf{f}', F, \mathbf{m}', M, \mathbf{t}', T)$. □

B. The Multidimensional Case[16]

Gender segregation has traditionally been associated with occupational segregation. However, a number of studies have shown that this one-dimensional approach is too restrictive: other job and worker characteristics, such as industry, private or public sector, ethnic group, level of education, job social status, and labour market status exhibit both trends and patterns of segregation which add to our understanding of occupational segregation.[17]

Thus, consider situations in which individuals can be classified in terms of a first characteristic, say educational attainment, indexed by $i = 1,\ldots, I$, and/or in terms of a second characteristic, say occupation, indexed by $j = 1,\ldots, J$.[18] Assume that there are J occupations in each category i, as well as I educational categories in each occupation j. As before, let F_{ij}, M_{ij}

by a change in W so that the relation between w_j and W changes but that between s_{fj} and s_{mj} does not.

[14] This is the case referred to as "a pair of one-way classification variables" in Mora and Ruiz-Castillo (2003a).

[15] Notice the analogy between this property and the additive decomposability property originally suggested in the income inequality literature by Bourguignon (1978) and Shorrocks (1980). For an alternative decomposition into three terms using the Gini-Segregation Index, see Silber (1989b), Boisso *et al.* (1994), Deutsch *et al.* (1994), and Sections 7.4 and 7.5 of Flückiger and Silber (1999). For the decomposition of the Karmel and MacLachlan segregation index into three terms see Borghans and Groot (1999).

[16] This is the case referred to as "a pair of two-ways classification variables" in Mora and Ruiz-Castillo (2003a).

[17] See, for instance, Jacobs (1989), Jacobsen (1994), Deutsch *et al.* (1994), Watts (1997), Blau *et al.* (1998), Blackburn *et al.* (2001), Charles (2003), and Mora and Ruiz-Castillo (2003a, 2003b, 2004).

[18] This paper only examines the case in which segregation takes places along two dimensions. However, the extension of these properties to more than two dimensions is straightforward. For an empirical study in which the non-student population of working age is classified according to human capital characteristics, labour market

and $T_{ij} = F_{ij} + M_{ij}$ be the number of females, males, and people, respectively, in occupation j in category i. Let $F_i = \Sigma_j F_{ij}$, $M_i = \Sigma_j M_{ij}$ and $T_i = \Sigma_j T_{ij}$ be the total number of females, males and people in category i, and let $\mathbf{f}^i = (F_{i1},..., F_{iJ})$, $\mathbf{m}^i = (M_{i1},..., M_{iJ})$, and $\mathbf{t}^i = (T_{i1},..., T_{iJ})$ be, respectively, the gender and people's frequencies across the J occupations in that category. Similarly, let $F_j = \Sigma_i F_{ij}$, $M_j = \Sigma_i M_{ij}$ and $T_j = \Sigma_i T_{ij}$ be the total number of females, males and people in occupation j, and let $\mathbf{f}^j = (F_{j1},..., F_{jI})$, $\mathbf{m}^j = (M_{j1},..., M_{jI})$, and $\mathbf{t}^j = (T_{j1},..., T_{jI})$ be, respectively, the gender and people's frequencies across the I educational categories in that occupation. Let $F = \Sigma_i F_i$, $M = \Sigma_i M_i$ and $T = \Sigma_i T_i$ be the overall number of females, males and people, respectively. Denote by **f**, **m**, and **t** the distributions of F, M, and T, respectively, across the I educational categories and J occupations of the economy. Finally, take $\theta^i(\mathbf{f}^i, F_i, \mathbf{m}^i, M_i, \mathbf{t}^i, T_i)$ and $\theta(\mathbf{f}', F, \mathbf{m}', M, \mathbf{t}', T)$ as measuring segregation within category i and between education characteristics, respectively, and define $\theta^j(\mathbf{f}^j, F_j, \mathbf{m}^j, M_j, \mathbf{t}^j, T_j)$ and $\theta(\mathbf{f}'', F, \mathbf{m}'', M, \mathbf{t}'', T)$ as measures of segregation within occupation j and between occupations, respectively. The following result is immediate:

Remark: (*Commutative Property*) If the segregation index θ satisfies A.14, then there exist υ_i and η_j with $\upsilon_i \geq 0$, $\eta_j \geq 0$ for each i and j, and $\Sigma_i \upsilon_i = \Sigma_j \eta_j = 1$, so that

$$\begin{aligned}\theta(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T) &= \Sigma_j \upsilon_j \theta^i(\mathbf{f}^i, F_i, \mathbf{m}^i, M_i, \mathbf{t}^i, T_i) + \theta(\mathbf{f}'', F, \mathbf{m}'', M, \mathbf{t}'', T)\\ &= \Sigma_j \eta_j \theta^j(\mathbf{f}^j, F_j, \mathbf{m}^j, M_j, \mathbf{t}^j, T_j) + \theta(\mathbf{f}', F, \mathbf{m}', M, \mathbf{t}', T).\end{aligned}$$

status, and occupations, see Mora and Ruiz-Castillo (2003b).

3. AN ENTROPY BASED INDEX OF SEGREGATION

3. 1. Definition and Motivation

In information theory, the expression

(1) $$I_j = w_j \log(w_j/W) + (1 - w_j) \log((1 - w_j)/(1 - W))$$

is known as the expected information of the message that transforms the proportions (W, (1 - W)) to a second set of proportions (w_j, (1 - w_j)). The value of this expected information is zero whenever the two sets of proportions are identical, it takes larger and larger positive values when the two sets are more different, and it is symmetrical in (w_j, (1 - w_j)). Therefore, I_j can be interpreted as an index of local segregation in occupation j within the approach reviewed in the previous section.

A weighted average of these J indices of local segregation will constitute an additive index of segregation. The selection of the weights is an important issue. One possible option is to give the same weight to each occupation, thus ensuring that the index is occupational invariant. However, we agree with England (1981) when she states: "The weighted index has more intuitive appeal. Suppose that occupations that segregate more (or less) grow faster over time, putting a greater (or lesser) number of persons into segregated work. I prefer an index that reveals this increase (or decrease) in segregation over one that adjusts the change out because it resulted from a change in the relative size of occupations that segregate to different extents." Thus, the I_E index of overall segregation is defined by

(2) $$I_E = \Sigma_j s_{tj} I_j.$$

That is to say, I_E is the weighted average of the information expectations, with weights proportional to the number of people in the occupations.[19]

The I_E index can also be motivated as an index of segregation which captures segregation whenever the frequency distributions of female and male workers across occupations differ from each other. To see this, note that from equation (2) it is straightforward to show that I_E can also be expressed as a weighted sum of two indexes:

$$I_E = W\ I_f + (1 - W)\ I_m$$

where $I_f = \Sigma_j\ s_{fj} \log(s_{fj}/s_{tj})$ and $I_m = \Sigma_j\ s_{mj} \log(s_{mj}/s_{tj})$. The choice of weights W and $(1 - W)$ ensures that the index I_E will give more weight to smaller deviations from $\{s_{tj}\}$ in the distribution across occupations of the majority gender.

3. 2. Basic Axioms

It is easily seen that I_E satisfies *Size Invariance* (A.1), that is to say, I_E is a relative index of segregation. The index I_E satisfies *Complete Integration* (A.2) because if $s_{fj} = s_{mj}$ for all j, then $s_{fj} = s_{tj}$ and $s_{mj} = s_{tj}$, so that $I_E = 0$. *Symmetry in Groups* (A.4), *Symmetry in Types* (A.5) and *Additivity* (A.13) follow directly from the definition of I_E.

I_E also fulfills *Complete Segregation* (A.3). Theil and Finizza (1971) show that I_E equals E -

19 See Mora and Ruiz-Castillo (2003a) for details on the seminal contribution to this approach by Theil and Finizza (1971) and Fuchs (1975). For a different segregation index also related to the concept of entropy, see

μ, where $E = W \log(1/W) + (1 - W)\log(1/(1 - W))$, $\mu = \Sigma_j s_{tj} E_j$, and $E_j = w_j \log(1/w_j) + (1 - w_j) \log(1/(1 - w_j))$.[20] Notice that E_j takes its minimum value, equal to 0, when $w_j = 0$. Otherwise, E_j is positive and reaches its maximum value, equal to log 2, when $w_j = 1/2$. To normalize E_j between 0 and 1, from here on it is assumed that all logarithms are in base 2. The same argument applies to E, which is also normalized to the unit interval. Now, if $w_j \in \{0,1\}$ for all j, then $E_j = 0$ for all j and $\mu = 0$, so that $I_E = E$. Given that μ is non-negative, I_E is bounded from above by E, which is itself bounded by 1. Therefore, I_E can only take values in the interval $[0, E] \subset [0, 1]$, and the index reaches its maximum when there is complete segregation.

To verify that I_E satisfies A.6 to A.10, it is useful to compute the marginal effect on I_E of an infinitesimal shift of the female population from occupation i to occupation k: $dF_k = - dF_i > 0$. From equation (2), we have that:

(3) $$dI_E = \left\{ \partial[T_k I_k]/\partial F_k - \partial[T_i I_i]/\partial F_i \right\} dF_k/T.$$

For any occupation j:

$$\partial[T_j I_j]/\partial F_j = I_j + T_j \left(\partial I_j/\partial w_j\right)\left(\partial w_j/\partial F_j\right),$$

where $\partial I_j/\partial F_j = \log(w_j/W) - \log((1-w_j)/(1-W))$ and $\partial w_j/\partial F_j = (1 - w_j)/T_j$, so that:

(4) $$\partial[T_j I_j]/\partial F_j = \log(w_j/W).$$

Hutchens (1991) and the discussion in Flückiger and Silber (1999).

[20] E and E_j are the entropy of a distribution with proportions (W, (1 – W)) and (w, (1 – w_j)), respectively. They measure the gender mix in the overall population and in occupation j, respectively.

Applying equation (4) to equation (3), it is seen after some manipulation that:

$$(5) \qquad dI_E = \log (w_k/w_i)\, dF_k/T.$$

For I_E, the *Weak Principle of Transfers* (A.6) follows directly from equation (5) and the fact that in a female dominated occupation, say i, $w_i > W$, whilst in a male dominated occupation, say k, $w_k < W$, so that $w_i > w_k$ and $dI_E < 0$. Of course, the decrease in the segregation index will take place as long as $w_i > w_k$, so the transfer does not have to occur between a female and a male dominated occupation.

To show that I_E satisfies *Movement between Groups* (A.7), note that given Equation (5), if $w_k > w_i$, then $dI_E > 0$ for a sufficiently small change $dF_k = - dF_i$. However, the condition for $dI_E > 0$, i.e. $w_k > w_i$, will always be met after any disequalizing change and, therefore, $dI_E > 0$ for any feasible discrete change, i.e. for any $0 < d \leq F_i$. Thus, A.7 is satisfied by index I_E. Since $w'_k > w_k > w_i > w'_i$, it is straightforward to see by a similar argument that I_E satisfies *Increasing Returns to Movement Between Groups* (A.10).

To show that I_E fulfils A.8, it is enough to show that if occupations i and k have equal gaps and $s_{ti} < s_{tk}$ then $dI_E < 0$. First, note that if i and k have equal gaps, then $T_i (w_i - W) = T_k (w_k - W)$. If $T_i < T_k$, then it follows that $w_i > w_k$. But then, from equation (5), $dI_E < 0$. Fulfillment of axiom A.9 directly follows from the fact that if $|s_{fi} - s_{mi}| > |s_{fk} - s_{mk}|$ and $T_i = T_k$, then $w_i >$

w_k. The proof that I_E satisfies *Zero Member Independence* (A.11) is immediate since $T_{J+1}/T = 0$.

A proof that I_E satisfies A.14 can be found in Mora and Ruiz-Castillo (2003a). On the other hand, *Insensitivity to Proportional Divisions* (A.12) holds because, as already stated, I_E satisfies both *Complete Integration* (A.2) and *Additive Decomposability* (A.14).

3. 3. Decomposability Properties

As already stated, a proof that I_E satisfies A.14 can be found in Mora and Ruiz-Castillo (2003a). This property is useful to attack the following classical problem. There is a potential bias due to small cell size (Blau *et al.*,1998): random allocations of individuals across occupations may generate high levels of gender segregation purely by chance. On the other hand, the use of more detailed categories leads to larger index values, since broader categories mask some of the segregation within them (England, 1981). Thus, it is interesting to study how far it is possible to aggregate an initial long list of occupations without reducing the gender segregation value too much. Herranz *et al.* (2005) propose an aggregation algorithm that uses I_E. The within-group term is identified as the error incurred in each step of the algorithm. Therefore, a reasonable stopping rule consists of selecting the furthest step for which the between group term is greater than or equal to the 1% bootstrapped lower bound for the original gender segregation value.

In the multidimensional case, both the decomposability property of I_E, as well as its commutative property, has been repeatedly used in a number of recent applications (see Mora and Ruiz-Castillo 2003a, 2003b, 2004). As an illustration, consider an economy in which people choose to work in an occupation either in the public sector A or in the private sector B. The population is said to be segregated in occupation j and sector i, i = A, B, whenever $w_j^i = F_j^i / T_j^i$ differs from $W = (F^A + F^B)/(T^A + T^B)$. The index I_E provides what is called a *direct* measure of gender segregation in occupation j and sector i in relation to the entire employed population:

$$I_E = \Sigma_i \Sigma_{j \in G_{ig}} s_{tj}\, s_j^i \{w_j^i \log (w_j^i / W) + (1 - w_j^i) \log ((1 - w_j^i)/(1 - W)\},$$

where $s_j^i = T_j^i / T_j$. This measure of overall gender segregation can be decomposed into a *between-group* term and a *within-group* term. First, consider the direct index of occupational segregation, that is, $I^B = \Sigma_j\, s_{tj}\, \{w_j \log(w_j/W) + (1 - w_j) \log((1 - w_j)/(1 - W))\}$. I^B can be interpreted as the *between-group* (direct) occupational gender segregation. On the other hand, the *within-group* gender segregation in the partition by occupations can be defined as $I^W = \Sigma_i s_t^i \Sigma_{j \in G_{ig}} s_{tj}^i \{w_j^i \log (w_j^i / w_j) + (1 - w_j^i) \log ((1 - w_j^i)/(1 - w_j))\}$, where $s_t^i = T^i / T$ and $s_{tj}^i = T_j^i / T^i$. As shown in Mora and Ruiz-Castillo (2003a), it turns out that $I_E = I^B + I^W$. This is a useful decomposition, where the term I^W measures the gender segregation induced by sector choice, the impact of occupational segregation being kept constant in I^B. Because of the commutative property discussed in Section 2.3, the index can also be decomposed into a term that captures the gender segregation induced by occupational choices within each sector, and a between-group term that captures the direct impact of sector choice on gender segregation.

4. CONCLUSIONS

This paper has reviewed some of the properties suggested in the methodological literature on the measurement of occupational gender segregation. It is found that an index of (relative) segregation based on the entropy concept, I_E, satisfies thirteen basic axioms previously proposed in the single-dimensional case, and can be expressed as the sum of a between-group and a within-group term both for any partition of the set of occupations and in the two-dimensional case. Moreover, it possesses an *Additive Decomposability* property (A.14) analogous to the one that serves to characterize the family of generalized entropy indices in the income inequality literature. Elsewhere, it has been shown that the index I_E can be interpreted as two different log-likelihood tests, so that bootstrap methods can be used to infer confidence intervals for small samples under general conditions and chi-square distributions can be used for large samples. It is also shown that both the *between* and the *within* elements in the decompositions can be understood as likelihood ratio tests and that the *within* term can be used to test differences in segregation across countries and over time (see Mora and Ruiz-Castillo, 2005).

How does I_E fare in relation to the remaining relative indices of gender segregation either widely used or recently suggested? Consider first indices that are not embedded in a statistical framework and restrict the attention to the single-dimensional case.

1. As pointed out in Zoloth (1976), James and Taeuber (1985), and Hutchens (1991), the well known Dissimilarity Index does not satisfy the strong versions of the Principle of Transfers, *Movement between Groups* (A.7) and *Increasing Returns to Movement between Groups* (A.10). A closely related index, originally suggested by Karmel and MacLachlan (1988), is

decomposable into 4 terms, one of which is margin-free. However, it does not satisfy A.7 and A.10 either, a fact that should be considered a serious drawback for a gender segregation index.

2. In the *marginal matching* (MM) approach advocated by Blackburn *et al.* (1993, 1995), occupational gender segregation is "the relationship between gendering of occupations and the sex of the workers, measuring the tendency for men and women to work in different occupations". MM was developed to measure changes over time in occupational segregation resulting from changes in the sex composition of occupations. However, MM does not satisfy *Movement between Groups* (A.7) and *Increasing Returns to Movement between Groups* (A.10). Finally, no proof for Additive Decomposability has been established.

3. The Gini segregation index satisfies (among other basic assumptions) A.4, A.7, and A.12. Although it violates A.13 and is not additively decomposable in the sense of A.14, it admits other decompositions (see footnote 14), and it remains an interesting index as demonstrated extensively in Flückiger and Silber (1999).

4. Hutchens (2001) characterizes a class of measures in terms of the axioms A.4, A.5, A.7, A.11, A.12, a version of A.13, and an invariance axiom not covered in this paper. In an important result, Hutchens (2004) went on to fully characterize a member of that class, called the square root segregation index,

$$\mathrm{H}(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T) = 1 - \Sigma_j \, (s_{fj} \, s_{mj})^{1/2},$$

in terms of the following eight axioms: A.2, A.3, A.4, A.5, A.7, A.12, A.14 and an invariance axiom. It can be shown that this index also satisfies A.1, A.6, A.7, A.8, A.9, A.10, A.12, and A.13. In brief, the square root index is the more comprehensive of all segregation indexes ever investigated and, therefore, deserves more applications than the only one that we know of with German data in Hutchens (2004).

As indicated before, none of the above indices has been embedded in a statistical framework, a property that has recently been emphasized in the following two cases.

5. The logarithmic index suggested by Charles and Grusky (1995),

$$A(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T) = \exp \left[(1/J) [\Sigma_j \ln (r_j) - (1/J) \Sigma_j \ln (r_j)]^2\right]^{1/2},$$

As pointed out in Watts (1998a, 1998b), this index does violate *Organizational Equivalence* (A.12). As indicated also by Watts (1998a, 1998b), the index is unduly influenced by extreme values caused by very low gender ratios that may characterize very small occupations. Moreover, if an occupation is completely segregated, with no (fe)male employees, the logarithm of the gender ratio $r_j = F_j/M_j$ is not defined.[21]

6. Like the I_E index advocated in this paper, Kakwani's (1994) preferred index

$$S_1(\mathbf{f}, F, \mathbf{m}, M, \mathbf{t}, T) = W(1 - W) \Sigma_j (s_{fj} - s_{mj})^2/s_{tj}$$

satisfies all basic axioms (except *Zero Member Independence*, A.11). Although it has not yet been attempted, it would appear that there exists a decomposition of S_1 involving a margin-free term. For his entire S_β family of indices, Kakwani (1994) defines a segregation index within a major occupation, $\theta^i(\mathbf{f}^i, F_i, \mathbf{m}^i, M_i, \mathbf{t}^i, T_i)$, and a between-group term, $\theta(\mathbf{f}', F, \mathbf{m}', M, \mathbf{t}', T)$, but it does not establish the additive decomposability in the sense of A.14. Nevertheless, the member S_1 of this family deserves more applications beyond the only one known to Australia that it is

[21] See, however, the reply by Grusky and Charles (1998).

contained in Kakwani (1994).

To sum up, in contrast to the entropy based index of segregation I_E studied in this paper, other existing measures of segregation either fail to satisfy one or more of the basic axioms in the single-dimensional case, do not possess decomposability properties, have not been motivated from a statistical approach, or (as shown in Mora and Ruiz-Castillo, 2005) are based on more restricted econometric models. However, the index I_E still awaits a full characterization in terms of a set of independent axioms.

REFERENCES

Anker, R. (1998), *Gender and jobs: Sex segregation of occupations in the world*. Geneva, ILO.

Blackburn, R.M., Jarman, J., Siltanen, J. (1993), "The Analysis of Occupational Gender Segregation over Time and Place: Considerations of Measurement and Some New Evidence", *Work, Employment and Society* **7**: 335-36.

Blackburn, R.M., Siltanen, J. and Jarman, J. (1995), "The Measurement of Occupational Gender Segregation: Current Problems and a New Approach", *Journal of the Royal Statistical Society A,* Part 2 **158**: 319-331.

Blackburn, R.M., Brooks B. and J. Jarman (2001), Occupational Stratification: The Vertical Dimension of Occupational Segregation, *Work, Employment & Society*, **15**: 511-38.

Blau, F., Simson, P. and Anderson, D. (1998), "Continuing Progress? Trends in Occupational Segregation Over the 1970s and 1980s", *Feminist Economics* **4**: 29-71.

Boisso, D., Hayes, K., Hirschberg, J. and Silber, J. (1994), "Occupational Segregation in the Multidimensional Case: Decomposition and Test of Significance", *Journal of Econometrics* **61**: 161-171.

Borghans, L., and Groot, L. (1999), "Educational Presorting and Occupational Segregation", *Labour Economics* **6**: 375-395.

Bourguignon, F. (1979), “Decomposable income inequality measures”, *Econometrica* **47**: 901-920.

Butler, R.J. (1987), "New Indices of Segregation" *Economic Letters* **24**:359-363.

Chakravarty, S.R. and Silber, J. (1992), "Employment Segregation Indices: An Axiomatic Characterization" In Eichhorn, W. (ed), *Models and Measurement of Welfare and Inequality*, New York: Springer-Verlag.

Charles, M. (1992), "Cross-National Variation in Occupational Sex Segregation", *American Sociological Review* **57**: 483-502.

Charles, M. (1998), "Structure, Culture, and Sex Segregation in Europe", *Research in Social Stratification and Mobility* **16**: 89-116.

Charles, M. (2003), "Deciphering Sex Segregation: Vertical and Horizontal Inequalities in Ten National Labor Markets", *Acta Sociologica* **46**: 267:87.

Charles, M. and Grusky, D. (1995), "Models for Describing the Underlying Structure of Sex Segregation", *American Journal of Sociology* **100**: 931-971.

Cortese, C.F., Falk, R.F., and Cohen, J.K. (1976), "Further Considerations on the Methodological Analysis of Segregation Indices", *American Sociological Review* **41**: 630-637.

Deutsch, J, Flückiger, Y. and Silber, J. (1994), "Measuring Occupational Segregation", *Journal of Econometrics* **61**: 133-146.

Duncan, O. and Duncan, B. (1955), "A Methodological Analysis of Segregation Indices", *American Sociological Review* **20**: 210-217.

England, P. (1981), "Assessing Trends in Occupational Sex Segregation, 1900-1976", in I. Berg (ed.), *Sociological Perspectives on Labor Markets*, New York, Academic Press.

Flückiger, Y. and Silber, J. (1999), *The Measurement of Segregation in the Labor Force*, Heidelberg, Physica-Verlag.

Fuchs, V. (1975), "A Note on Sex Segregation in Professional Occupations," *Explorations in Economic Research* **2**: 105-111.

Grusky, D.B. and Charles, M. (1998), "The Past, Present, and Future of Sex Segregation Methodology", *Demography* **35**: 497-504.

Herranz, N., Mora, R., and Ruiz-Castillo, J. (2005), "An Algorithm to Reduce the Occupational Space in Gender Segregation Studies", *Journal of Applied Econometrics*. **20**: 25-37

Hutchens, R. M. (1991), "Segregation Curves, Lorenz Curves and Inequality in the

Distribution of People Across Occupations", *Mathematical Social Sciences* **21**: 31-51.

Hutchens, R. M. (2001), "Numerical Measures of Segregation: Desirable Properties and Their Implications", *Mathematical Social Sciences* **42**: 13-29.

Hutchens, R. M. (2004), "One Measure of Segregation", *International Economic Review*. **45**: 555-578.

Jacobs, J. (1989), "Long-Term Trends in Occupational Segregation by Sex", *American Journal of Sociology* **95**:160-173.

Jacobsen, J. (1994), "Trends in Work Force Sex Segregation, 1960-1990", *Social Science Quarterly* **75**: 204-211.

James, D.R. and Taeuber, K.E. (1985), "Measures of Segregation", in G. Schmid and R. Weitzel (eds.), *Sex Discrimination and Equal Opportunity: The Labor Market and Employment Policy*, London, Gower Publishing Company.

Jonung, C. (1984), "Patterns of Occupational Segregation by Sex in the Labor Market", in N.B. Tuma (ed.), *Sociological Methodology*, San Francisco, Jossey-Bass.

Kakwani, N.C. (1994), "Segregation by Sex: Measurement and Hypothesis Testing", *Research on Economic Inequality* **5**: 1-26.

Karmel, T. and MacLachlan (1988), M., "Occupational Sex Segregation: Increasing or Decreasing?", *Economic Record* **64**:187-195.

Kolm, S.C. (1999), "The Rational Foundations of Income Inequality Measures", in J. Silber (ed.), *Handbook of Income Inequality Measurement*, Dordretch, Kluwer Academic Publishers.

Lewis, D.E. (1982), "The Measurement of the Occupational and Industrial Segregation of Women", *Journal of Industrial Relations* **24**: 406-423.

Moir, H., and Selby Smith, J. (1979), "Industrial Segregation in the Australian Labour Market", *Journal of Industrial Relations* **21**: 281-291.

Mora, R. and Ruiz-Castillo, J. (2003a), "Additively Decomposable Segregation Indices. The Case of Gender Segregation By Occupations and Human Capital Levels In Spain", *Journal of Economic Inequality* **1**: 147-179.

Mora, R. and Ruiz-Castillo, J. (2003b), "Gender Segregation: From Birth to Occupation", Working Paper No. 03-36, Economic Series 12, Universidad Carlos III, Madrid.

Mora, R. and Ruiz-Castillo, J. (2004), "Gender Segregation by Occupations in the Public and the Private Sectors. The Case of Spain In 1977 and 1992", *Investigaciones Económicas* **XXVIII**: 399-428.

Mora, R. and Ruiz-Castillo, J. (2005), "An Evaluation of an Entropy Based Index of Segregation", *mimeo*, Universidad Carlos III Madrid.

Peach, C. (1975), *Urban Residential Segregation*, London: Longman.

Schwartz, J. and Winship, C. (1979), "The Welfare Approach to Measuring Inequality", in K.F. Schuessler (ed.), *Sociological Methodology*, San Francisco: Jossey-Bass.

Shorrocks, A. F. (1980), "The Class of Additively Decomposable Inequality Measures", *Econometrica* **48**: 613-625.

Shorrocks, A. F. and J. Foster (1987), "Transfer Sensitive Inequality Measures", *Review of Economic Studies* **54**: 485-497.

Silber, J. (1989a), "On the Measurement of Employment Segregation", *Economic Letters* **30**: 237-242.

Silber, J. (1989b), "Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality", *Review of Economics and Statistics* **LXXI**: 107-115.

Silber, J. (1992), "Occupational Segregation Indices in the Multidimensional Case: A Note", *Economic Record* **68**: 276-277.

Siltanen, J. (1990), "Social Change and the Measurement of Occupational Segregation by Sex: An Assessment of the Sex-Ratio Index", *Work, Employment and Society* **4**: 1-29.

Theil, H. and Finizza, A.J. (1971), "A Note on the Measurement of Racial Integration of Schools by Means of Information Concepts", *Journal of Mathematical Sociology* **1**: 187-194.

Watts, M. (1992), "How Should Occupational Segregation Be Measured?", *Work, Employment and Society* **6**: 475-487.

Watts, M. (1997), "Multidimensional Indices of Occupational Segregation", *Evaluation Review* **21**: 461-482.

Watts, M. (1998a), "Occupational Gender Segregation: Index Measurement and Econometric Modelling", *Demography* **35**: 489-496.

Watts, M. (1998b), "The Analysis of Sex Segregation: When is Index Measurement Not Index Measurement?", *Demography* **35**: 505-508.

Zoloth, B.S. (1976), " Alternative Measures of School Segregation", *Land Economics* **52**: 278-298.